multiple levels of parallelism offered by modern CPUs : Thread-level parallelism (TLP) is provided by expanding the particle transport physics to multiple CPU cores. On this level, the MC simulation is embarrassingly parallel so that each core runs its own simulation. However, gathering energy deposition events into one dose cube is challenging and needs to be orchestrated carefully using a buffered memory streaming technique. Data-level parallelism is achieved be re-engineering vital parts of the engine for Singe Instruction Multiple Data (SIMD) registers. Portability between a wide range of CPU generations is achieved by using APIs such as OpenMP, Cilk Plus and Intel MKL. PhiMC includes a source model for the dose delivery system and is capable of simulating clinical IMRT cases in the required spatial resolution of the dose grid. PhiMC has been implemented in C++ and the accuracy of the algorithms is validated against the original DPM framework using the Two One-Sided statistical Test (TOST) procedure.

Results: We could show that the dose values calculated with PhiMC and DPM coincide at levels of 100% (for electrons) and 99.7% (for photons) with a significance level of 0.05. 900 million particle histories in water sampled from a 6 MV photon spectrum can be simulated in 47.3 seconds with a mean uncertainty of 0.32% on a dual Xeon workstation. An IMRT prostate plan with 9 beams and a voxel resolution of (1.95 x 1.95 x 2.0) mm^3 could be simulated with 150 million histories in 16.2 seconds achieving a mean statistical uncertainty of 0.86%. A head&neck patient case using 9 beams and a voxel resolution of (1.56 x 1.56 x 3.0) mm^3 can be simulated in 8.2 seconds with an uncertainty of 0.89%.

Conclusions: The dose of typical clinical IMRT plans can be simulated in less than 20 seconds with a mean uncertainty of less than 1% with our CPU-based PhiMC dose calculation framework. Compared to a GPU-based implementation of DPM (gDPM), PhiMC achieves better runtime results. In addition to that, a CPU implementation is not limited to a few GB of RAM and can therefore simulate larger plans or plans with a higher resolution.

PO-0885

Validation of a head and neck DVH prediction model for use in commercially available knowledge-based planning software

L. Olsen[1], J. Kavanaugh[1], M. Yang[1], H. Gay[1], W. Thorstad[1], S. Mutic[1]

[1]Washington University School of Medicine, Radiation Oncology, Saint Louis, USA

Purpose/Objective: A knowledge-based planning (KBP) algorithm was recently implemented into a commercial treatment planning system. The objective of this work was to evaluate the impact of training set plan quality on the ability of the KBP software to produce a high quality head and neck (HN) DVH prediction model and to validate the KBP model for clinical use.

Materials and Methods: The KBP software was used to create an initial model for HN patients using a training set of 149 plans from our clinical database. The training set included bilateral/unilateral, definitive and post-operative HN cases with arange of simultaneously integrated boost prescription dose levels determined from the clinical diagnosis. DVH prediction models were trained for the brain, brainstem, esophagus, larynx, lips, oral cavity, parotid glands, pharyngeal constrictors, spinal cord, and submandibular glands. Optic structures were excluded from the modeling process due to small structure volume and standard maximum dose objectives were used for optimization. The quality of each plan in the training cohort was assessed by comparing the clinically-approved DVHs to the DVHs predicted by the initial model. This method was used to identify dosimetric outliers where the clinical DVHs were sub-optimal to the predicted DVHs. Identified outliers (Table 1) were removed from the training set to create a refined model. A randomly selected 20 patient validation cohort was used to evaluate the impact of training set plan quality on the output of the DVH prediction model. This impact was assessed by comparing clinically relevant DVH metrics (mean dose or maximum dose) predicted by the refined model to those predicted by the initial model. The accuracy of the refined model to achieve predicted DVH values was evaluated by comparing DVH metrics predicted by the refined model to DVH metrics achieved in a replan using the refined HN model. Clinical acceptability was evaluated by comparing the DVH metrics in the clinically approved plan to the KBP model replan.

Results: There was a a reduction in the average predicted mean/maximum dose in the refined model compared to the initial model (Table 1) for all structures included in the model. The refined HN model demonstrated excellent predictive accuracy, with a small average difference between clinical metrics of the KBP modal replan and the refined DVH predictions (Table 1). The ability of the HN model to produce clinically acceptable plans was demonstrated in the 20 validation patients by similar clinical metrics achieved between the clinical plan and the KBP replan.

Structure	Parotid				Submandibular			
# Outliers	82				98			
DVH Metric	Dmean (Gy)				Dmean (Gy)			
Plan	Clinical	Initial Pred.	Refined Pred.	KBP Replan	Clinical	Initial Pred.	Refined Pred.	KBP Replan
Average	31.36	32.07	30.24	30.66	54.21	52.76	51.63	52.78
St Dev	21.23	19.42	18.99	20.73	21.21	17.20	17.39	20.52
Structure	Lips				Pharyngeal Constrictor			
# Outliers	26				98			
DVH Metric	Dmean (Gy)				Dmean (Gy)			
Plan	Clinical	Initial Pred.	Refined Pred.	KBP Replan	Clinical	Initial Pred.	Refined Pred.	KBP Replan
Average	18.57	21.54	17.74	16.94	50.36	50.56	49.24	49.55
St Dev	9.58	8.47	8.04	9.04	15.44	14.14	14.51	15.64
Structure	Larynx				Oral Cavity			
# Outliers	60				40			
DVH Metric	Dmean (Gy)				Dmean (Gy)			
Plan	Clinical	Initial Pred.	Refined Pred.	KBP Replan	Clinical	Initial Pred.	Refined Pred.	KBP Replan
Average	45.88	47.78	43.02	44.59	44.59	45.58	44.59	44.53
St Dev	17.75	11.40	14.57	17.01	18.98	17.34	17.83	18.84
Structure	Brainstem				Spinal Cord			
# Outliers	55				30			
DVH Metric	Dmean (Gy)				Dmax (Gy)			
Plan	Clinical	Initial Pred.	Refined Pred.	KBP Replan	Clinical	Initial Pred.	Refined Pred.	KBP Replan
Average	10.19	12.77	12.43	10.39	37.45	38.24	36.62	38.93
St Dev	7.17	7.90	7.30	7.02	2.01	2.59	2.72	2.76
Structure	Esophagus				Brain			
# Outliers	55				34			
DVH Metric	Dmean (Gy)				Dmean (Gy)			
Plan	Clinical	Initial Pred.	Refined Pred.	KBP Replan	Clinical	Initial Pred.	Refined Pred.	KBP Replan
Average	33.66	35.20	31.65	31.30	3.15	3.42	3.50	2.97
St Dev	14.22	10.84	12.46	14.13	2.36	2.38	2.49	2.21

Conclusions: This study validates that the developed KBP model systematically produces clinically acceptable plans for HN patients. The results demonstrate the impact of training

set plan quality on high-quality output of the KBP software. The validated HN model will be available for use in an upcoming clinical release of the KBP software.

PO-0886

Three phase adaptive 18F-FDG-PET-voxel intensity-based VMAT versus 6-beam IMRT for head-and-neck cancer

A.M.L. Olteanu[1], B. Speleers[2], D. Berwouts[1], W. De Neve[1], F. Duprez[1], T. Vercauteren[1], W. De Gersem[2]
[1]*UZ Ghent, Radiotherapy, Ghent, Belgium*
[2]*Ghent University, Radiotherapy, Ghent, Belgium*

Purpose/Objective: In this study we investigated the implementation of a new class-solution of volumetric modulated arc therapy (VMAT) for a three-phase adaptive ^{18}F-FDG-PET-voxel-based dose-painting-by-numbers (DPBN) dose-escalation treatment. VMAT dose distributions were compared to the ones made using a standard 6-beam static intensity modulated radiotherapy (sIMRT) technique.

Materials and Methods: 10 non-metastatic head-and neck cancer patients, enrolled in the adaptive arm of a phase II DPBN trial were planned both with sIMRT and VMAT. Separate treatment plans based on two pre and per-treatment (after the 8th fraction) ^{18}F-FDG-PET/CTs and one per-treatment CT (after the 18th fraction) were made according to the trial protocol (Table 1). Dose distributions were summed on the pretreatment CT. Plans were evaluated in terms of dose levels, dose painting quality factors (QFs), treatment time and verified with Delta^4 (Scandidos, Uppsala, Sweden) measurements.

Table 1. Prescribed fraction dose for the three phase adaptive treatment protocol. The treatment plans for the first 2 phases were based on the ^{18}F-FDG-PET/CT information, while for the third phase only CT data was used.

Abbreviations: GTV = gross tumor volume; PTV_{HR} = high-risk planning target volume (3 mm expansion of the high-risk clinical target volume - CTV_{HR}); CTV_{HR} = a three-dimensional GTV expansion of 1 cm adjusted to air cavities and uninvolved bones.

	Phase I	Phase II	Phase III
GTV (dose painting range)	2.2 – 3.1 Gy	2.2 – 3.1 Gy	
PTV_{HR}	2.0 Gy	2.0 Gy	2.2 Gy

Results: VMAT plans allowed the same level of dose escalation in the targets, while significantly reducing the dose to organs-at-risk (OARs). On average, the percentage of the ipsilateral parotid volume receiving at least 27 Gy was reduced from 44.0% to 38.8% and its median dose from 22.8 to 19.5 Gy (p<0.05). Gross tumor volume QFs were significantly improved with VMAT. In 17 out of 20 phase I and phase II treatment plans, VMAT QF was better (maximum improvement 2.1%), while for the rest it was similar to sIMRT. Planning time of both techniques was similar and arc treatment delivery was 2 to 3 times faster. The Delta 4 measurements were in very good agreement with the dose calculation for both types of plans.

Conclusions: Biologically-guided volumetric modulated arc therapy is able to increase the sparing of OARs compared to sIMRT without compromising target doses or treatment delivery quality. Thus, it becomes a valuable technique for an adaptive treatment strategy, which follows the anatomical patient changes through the treatment time. The significantly faster VMAT delivery reduces the risk on intra-fraction movement.

PO-0887

Fast and realistic Monte Carlo evaluation of the robustness of proton therapy plans

K. Souris[1], J.A. Lee[1], E. Sterpin[1]
[1]*UCL - IREC, Molecular Imaging Radiology and Oncology (MIRO), Bruxelles, Belgium*

Purpose/Objective: Proton range uncertainties jeopardize the theoretical advantage of intensity modulated proton therapy over photon-based modalities. Depending on the heterogeneity level, this range uncertainty can grow up to 4.6% of the nominal range + 1.2mm with typical analytical dose calculation methods (Paganetti 2012, PMB). The robustness of treatment plans can be further evaluated by simulating possible realizations of uncertainties with systematic and random components. Such a strategy may be a daunting task for analytical algorithms because computation time scales linearly with the number of scenarios simulated, which increases strongly if random errors are considered. Monte Carlo (MC) simulations offer here a double advantage: 1) they reduce the range uncertainty down to 2.4% + 1.2 mm; 2) they potentially allow simulating random errors with no significant increase in computation time. This study employs a fast MC tool, which can compute the impact of random errors in a single simulation.

Materials and Methods: MCsquare, the new software created for this study, implements optimized algorithms on the Xeon Phi coprocessor to accelerate MC computations. MCsquare can compute a dose distribution in less than one minute. Multiple uncertainty scenarios are created. A 2.5 mm systematic setup error is modeled by shifting the CT image in all 6 directions of space. Random setup errors are modeled by a 1 mm random shift for each particle simulated by the MC engine. The uncertainty in the conversion from Hounsfield units to stopping powers is taken into account by applying a +/- 3% uniform bias to the patient densities. The experiment involves a water phantom, considering both a traditional plan with a PTV (2.5mm isotropic margin) and a robust plan (3% density uncertainty, 2.5 mm systematic setup errors). The CTV surrounds a circular organ-at-risk. The random error model employed in this study considers a large number of sampling, meaning an infinite number of fractions. The second experiment aims at determining the minimal number of fractions required to ensure the validity of this approximation. For this purpose, various sequences of fractions are generated, with different random errors.

Results: The robustness of the treatment plan is easily verified by looking at the deviations of the DVH curve with respect to the nominal plan (red curve). The robust plan shows small deviations compared to the traditional PTV plan. Considering only random errors, the DVH distributions no longer vary for treatment with more than 30 fractions. This result validates the assumption of the infinite random sampling for our robustness test for typical fractionation strategies.